Appl. No.:   09/418,418
Amdt. dated July 18, 2003
Reply to Office action of March 26, 2003

REMARKS

Claims 1-12 and 14-21 are currently pending in this case. Claims 1-12 and 14-21 have been rejected on various grounds as discussed below. The Applicants respectfully traverse the rejections. Reconsideration of Claims 1-12 and 14-21 is respectfully requested in view of the following remarks.

The independent Claims 1, 20 and 21 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over a publication by Chakrabarti et al. (Automatic Resource Compilation by Analyzing Hyperlink Structure and Associated Text, April 14, 1998). In particular, the Examiner asserts that Chakrabarti teaches the step of "determining which of the hypertext documents are expert documents", which, among other limitations, is required in Claims 1-12 and 14-21. The Applicants submit that Chakrabarti does not form a set of expert documents from the set of all documents crawled. Instead, each step of Chakrabarti ranks documents based on the topic of a query and forms sets of documents related to the topic. In contrast, the claimed invention requires identifying expert documents from the entire set of documents without limitation by a particular topic. Only after the claimed invention has identified the set of experts, does it use a query to identify experts relevant to a particular topic from the set created in the first step and expands the list to include target documents to which the relevant experts point.

In the present specification at page 4, lines 4-7, it is stated that: "Expert pages are preferably identified in a pre-processing step in which a subset of the pages crawled by the search engine are identified as experts (for example, 2.5 million of 140 million pages may be identified as experts)." Such a pre-processing step may be done for all pages crawled by the search engine, e.g. 140 million pages. It may be done before a query is actually received, since it does not use a topic in the first step. The set of experts is not limited to any particular topic submitted in any particular query. This is also stated on Page 8, lines 25-27 as follows: "Step 202 preferably processes a search engine's database of pages and selects a subset which is considered to be good sources of links on specific topics, albeit unknown." The expert list is created by

measuring the out-degree of each page and checking to be sure the page points to non-affiliated hosts, see for example Fig. 3(a). In Fig. 3(b) the experts may be further limited to those which point to pages with the same broad classification, but this is not limited to any particular topic. Since the expert set creation step does not depend on having received a query with a topic, it can be performed occasionally, for example at off peak demand times, and stored for use when queries are received.

The second step of the present claims is to rank the expert documents in accordance with a current search query. This is the first step in which the topic of the query becomes part of the process. The preselected expert list may be searched for pages relevant to the topic, not the entire set of pages. That is, it is not necessary to search the entire set of pages crawled by the search engine. In the above example, only 2.5 million pages needs to be searched, not 140 million. The present invention then uses the ranked expert pages which are relevant to the topic to expand the list of target documents by ranking the pages to which the relevant experts point. It is not necessary to search the entire web to find documents relevant to the topic, because the ranked expert documents point to them.

In contrast, the process of Chakrabarti begins with a topic from a query and searches the entire set of pages accessible to the search engine for pages relevant to the topic. For example in Section 2. Algorithm, second paragraph, Chakrabarti states: "Given a topic, the algorithm first gathers a collection of pages from among which it will distill ones that it considers the best for the topic. ... The topic is sent to a term-based search engine – *AltaVista* in our case – and a root set of 200 documents containing the topic term(s) is collected. ... The root set is then augmented through the following *expansion* step: we add to the root set (1) any document that points to a document in the root set, and (2) any document that is pointed to by a document in the root set. We perform this expansion step twice..." To the extent that Chakrabarti expands by identifying documents that point to the root set, it is searching the entire web again.

Appl. No.:   09/418,418
Amdt. dated July 18, 2003
Reply to Office action of March 26, 2003

Chakrabarti then uses hub scores and authority scores to rank the documents in the augmented set. In Section 2. Algorithm, fourth paragraph, Chakrabarti discusses generation of a hub score h(p) and an authority score a(p) for each page in the augmented set. In Section 2. Algorithm, fifth paragraph, Chakrabarti discusses weighting the hub and authority scores based on topic references as explained in more detail in Section 2.1. In Section 2. Algorithm, sixth paragraph, Chakrabarti computes vectors of the weighted hub and authority scores and selects the pages with the top fifteen scores for each of hub and authority values. Chakrabarti does not actually discuss experts, much less identifying a set of experts without reference to a particular topic submitted in a query.

In view of the above remarks, the Applicants submit that Claims 1, 20 and 21 are clearly patentable over the Chakrabarti reference. Since the remaining claims all depend from Claims 1, 20 and 21, the remaining claims should also be patentable over the Chakrabarti reference. Applicants respectfully request reconsideration and allowance of Claims 1-12 and 14-21.

If the Examiner feels that a telephone conference would expedite the resolution of this case, he is respectfully requested to contact the undersigned.
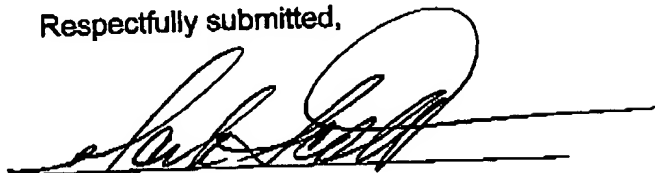
In the course of the foregoing discussions, Applicants may have at times referred to claim limitations in shorthand fashion, or may have focused on a particular claim element. This discussion should not be interpreted to mean that the other limitations can be ignored or dismissed. The claims must be viewed as a whole, and each limitation of the claims must be considered when determining the patentability of the claims. Moreover, it should be understood that there may be other distinctions between the claims and the prior art which have yet to be raised, but which may be raised in the future.

If any fees or time extensions are inadvertently omitted or if any fees have been overpaid, please appropriately charge or credit those fees to Hewlett-

Appl. No.:　09/418,418
Amdt. dated July 18, 2003
Reply to Office action of March 26, 2003

Packard Company Deposit Account Number 08-2025 and enter any time extension(s) necessary to prevent this case from being abandoned.

Respectfully submitted,

Mark E. Scott
PTO Reg. No. 43,100
CONLEY ROSE, P.C.
(713) 238-8000 (Phone)
(713) 238-8008 (Fax)
ATTORNEY FOR APPLICANTS

HEWLETT-PACKARD COMPANY
Intellectual Property Administration
Legal Dept., M/S 35
P.O. Box 272400
Fort Collins, CO　80527-2400

104102.01/2162.14900